三种厚朴叶绿体基因组的比较研究

张敏[1]，尹彦棚[1]，周罗静[1]，任波[1]，王丽[2]，时小东[3]，侯飞侠[1]，彭成[1]，高继海[1*]

（1. 成都中医药大学 药学院 西南特色中药资源国家重点实验室，成都 611137；2. 四川省林业科学研究院，成都 610065；3. 成都大学，成都 610106）

摘　要：为了深入发掘日本厚朴、厚朴、凹叶厚朴叶绿体基因组差异，筛选厚朴优良性状候选基因，开展三种厚朴的分子遗传研究，该文利用 Illumina HiSeq 高通量测序平台首次对日本厚朴叶绿体进行测序、组装，并与已有的厚朴、凹叶厚朴叶绿体基因组共同注释，获得三个物种叶绿体基因图谱，筛选出三个基因组中的差异基因，又与同科中 11 个亲缘物种进行叶绿体基因组比对，构建 NJ 遗传树。结果表明：日本厚朴叶绿体基因组的 Clean Reads 为 19 791 019，Q30 为 91.33%，组装后基因组全长 160 051 bp，GC 含量为 39.2%，含 tRNA 37 个，rRNA 8 个。比对分析发现三种厚朴具有相似的 IR、LSC 和 SSC 结构，以及 GC 含量和 tRNA 数量，但编码基因种类和数量、内含子和外显子的数量和结构等存在差异。日本厚朴的功能基因数目较厚朴、凹叶厚朴分别多 6 个和 4 个，主要分布于 LSC 区和 IR 区，涉及核糖体大亚基、核糖体小亚基和未知功能基因类群。系统发育分析结果进一步显示日本厚朴与凹叶厚朴亲缘关系较近，其次是厚朴。该研究表明日本厚朴具有更丰富的叶绿体基因组结构、组成和变异特征，是其适应高纬度地区弱光、低温环境的分子机制，这为厚朴类优良品种的分子选育提供有力的指导。

关键词：厚朴，凹叶厚朴，日本厚朴，叶绿体基因组，系统发育树

中图分类号：Q943　　　　**文献标识码**：A

Comparative study on chloroplast genomes of the three *Magnolia* species

ZHANG Min[1], YIN Yanpeng[1], ZHOU Luojing[1], REN Bo[1], WANG Li[2], SHI Xiaodong[3], HOU Feixia[1], PENG Cheng[1], GAO Jihai[1*]

(1. *Pharmacy College, Chengdu University of Traditional Chinese Medicine, Key Laboratory of Distinctive Chinese Medicine Resources in Southwest China*, Chengdu 611137, China; 2. *Sichuan Academy of Forestry Sciences*, Chengdu 610065, China; 3. *Chengdu University*, Chengdu 610106, China)

Abstract: In order to investigate the good genes, cultivate the main superior cultivars and discover phylogenetic relationships of *Magnolia officinalis*, *M. officinalis* subsp. *biloba* and *M. hypoleuca*, we compared the differences among the cp (chloroplast) genomes of three *Magnolia* species and performed a phylogenetic tree of 14 species. Illumina HiSeq platform was used to sequence and assemble the cp genome of *M. hypoleuca*. Then the cp genomes

---

基金项目：“杏林学者”学科人才提升计划（QNXZ2018017，QNXZ2019001）；四川省首批中医药学科建设重点项目（药用植物学，川中医药函[2020]84 号）；西南特色中药资源基因组学创新平台（2020ZYD058）[Supported by “Xinglin Scholar” Discipline Talents Promotion Program (QNXZ2018017，QNXZ2019001); The First Key Project of Traditional Chinese Medicine Discipline Construction in Sichuan Province (Medicinal Botany, Sichuan Medical Han [2020] No. 84); Genomics Innovation Platform of Chinese Medicine Resources with Southwestern Characteristics (2020ZYD058)]。

作者简介：张敏（2002-），从事中药资源与开发、种质资源研究，（E-mail）515638246@qq.com。

*通信作者：高继海，博士，副教授，从事中药资源与开发、种质资源研究，（E-mail）gaojihai@cdutcm.edu.cn。

of three *Magnolia* species were annotated by online platform and performed with three *Magnolia* species cp gene cycles. Moreover, the cp genomes of other 11 *Magnolia* species were downloaded from the NCBI database and phylogenetic tree of 14 all species cp genomes was constructed based on NJ method. The results indicated that Clean Reads of *M. hypoleuca* were 19 791 019, and Q30 was 91.33%. The total length of cp genome of *M. hypoleuca* was 160 051 bp, its GC content was 39.2%, including 37 tRNA and 8 rRNA. Besides, there were respectively six and four more functional gene numbers of *M. hypoleuca* than the other two *Magnolia* species, which indicated it had stronger viability. Compared with the cp genome structure of three *Magnolia* species, it showed that three *Magnolia* species had similar IR, LSC and SSC structures, GC content and tRNA number, but there were differences in the type and number of coding genes, the number and structure of introns and exons. Besides, the differential functional genes of three *Magnolia* species were mainly located in LSC region and IR region, involving large ribosomal subunits, small ribosomal subunits and unknown functional genes groups. Furthermore, based on NJ phylogenetic tree, *M. hypoleuca* was closely related to *M. officinalis* subsp. *biloba*, next to *M. officinalis*. In this study, *M. hypoleuca* had more abundant cp genome structure, composition and variation characteristics, which was the molecular mechanism of its adaptation to low light and low temperature environment in high latitude area. And it will also provide strong guidance for molecular breeding of excellent *Magnolia* varieties.

Key words: *Magnolia officinalis*, *Magnolia officinalis* subsp. *biloba*, *Magnolia hypoleuca*, chloroplast genome, phylogenetic tree

厚朴（*Magnolia officinalis*）、凹叶厚朴（*M. officinalis* subsp. *biloba*）及日本厚朴（*M. hypoleuca*）均为木兰科药、材两用经济林阔叶树种，广泛分布于韩国、日本和中国等地区，具有水土保持、绿化、美化环境和药用等多种应用价值（彭梦婕，2020）。目前三种厚朴的相关研究主要集中在化学成分、临床应用及资源调查（胥爱丽等，2021）等方面。然而，日本厚朴非我国原产资源，生长于相对低温环境中，具备许多优于另两者的生物学性状，如其耐寒能力相比于其他木兰属植物更强（Kwon & Oh, 2015），生长速度和成熟率都比同类植物更快，尤其是在生长幼期，以每年 60~90 cm 的速度成长（Oguchi et al., 2017），是解析和培育厚朴类物种优良性状的理想材料。然而，日本厚朴的分子遗传学研究资料匮乏，且三种厚朴的生长速度和抗寒能力差异的形成原因缺乏分子生物学机制研究，进而影响了厚朴类优良品种的选育。

叶绿体是植物细胞中必不可少的细胞器，在光合作用、固碳以及氨基酸的合成中起到重要作用，其全基因组包含大量遗传信息，结构高度保守，被广泛应用于植物分子进化及系统发育的研究，并在药用植物的遗传转化、基因工程和分子选育等方面发挥着重要作用（赵祺等，2021）。目前，厚朴和凹叶厚朴已有叶绿体基因组报道，而具有更多生物特性的日本厚朴却缺乏叶绿体基因组的研究。鉴于此，本研究利用 Illumina HiSeq 高通量测序技术，对日本厚朴进行了叶绿体基因组测序，与既有的厚朴、凹叶厚朴叶绿体基因组比对，筛选出了诸多差异的核苷酸序列，破译三种厚朴之间的亲缘关系，找出三种厚朴具有更强生长发育能力和耐寒性等功能的相关基因，在一定程度上为以更强的生长速度和抗寒能力为优良性状的目标物种提供了依据。

1 材料与方法

1.1 材料及预处理

日本厚朴新鲜、幼嫩、健康的叶片于 2019 年 10 月 10 日采自四川省成都市温江区成都中医药大学药用植物园（30°42′ E，103°49′ N），经国家中药种质资源库专家高继海副教授鉴定所有样品均来源于木兰科木兰亚属日本厚朴（*M. hypoleuca*）。三种厚朴的命名参考于中国植物物种信息数据库（http://db.kib.ac.cn/）。叶片

使用无菌水擦拭干净，迅速冻存备用。凭证标本（馆藏序列号为 ZYC190910）保存于成都中医药大学中医药传统文化博物馆。

叶绿体提取步骤如下：新鲜叶片于液氮中研磨后悬于一定量 A 液（50 mmol·L^{-1} Tris，25 mmol·L^{-1} EDTA，1.25 mol·L^{-1} NaCl，0.25 mmol·L^{-1} Vc，1.5% PVP，pH 3.6）中，4 层纱布过滤，4 ℃下 200 *g* 离心 20 min，取上清；加入常温缓冲液 B（50 mmol·L^{-1} Tris, 25 mmol·L^{-1} EDTA, 1.25 mol·L^{-1} NaCl, 0.25 mmol·L^{-1} Vc, 1 mmol·L^{-1} DTT, 0.1%牛血清蛋白 BSA, pH 8.0），常温静置，4 ℃下 2 000 *g* 离心 10 min，弃上清，叶绿体沉淀于 4 ℃保存备用。

1.2 叶绿体基因组提取和测序

针对叶绿体材料，采用改良的 CTAB 法（Matthes et al., 2020）分离提取 DNA。DNA 经检测合格后，用超声波机械打断，再进行片段纯化、末端修复、3′端加 A、连接测序接头，进行 PCR 扩增形成测序文库，最后使用高通量测序平台 Illumina HiSeq PE150 进行测序。测序得到的原始测序序列（Raw Reads），里面含有带接头的、低质量的 Reads，为了保证信息分析质量，对 Raw Reads 进行过滤、质控，得到 Clean Reads，用于后续信息分析。数据过滤的主要步骤如下：（1）去除带接头的 Reads；（2）过滤 N 含量超过 10%的 Reads；（3）去除质量值低于 10 的碱基超过 50%的 Reads。对过滤后的高质量数据随机抽取 2 000 条 Reads 数据，通过 Blast 软件比对 NT 库检测样品是否受到污染。

1.3 叶绿体基因组的组装和注释

原始序列上传于国家生物信息中心数据库（序列号 PRJCA004348）。利用 Galaxy 在线平台(https://usegalaxy.org)对日本厚朴叶绿体基因组测序结果进行组装(Yan et al., 2015)，且下载厚朴(*M. officinalis*, NC_020316）和凹叶厚朴（*M. officinalis* subsp. *biloba*, JN867581）的叶绿体基因组 FASTA 文件，合并后作为日本厚朴的参考基因文件。再通过 CPGAVAS2 在线平台(http://www.herbalgenomics.org/cpgavas2)完成厚朴、凹叶厚朴、日本厚朴叶绿体基因组的注释。

1.4 聚类分析

除厚朴、凹叶厚朴外，又于 NCBI 数据库中下载荷花玉兰（*M. grandiflora*, JN867584）、星花玉兰（*Yulania stellata*, NC_039941）、天女木兰（*Oyama sieboldii*, NC_041435）、望春玉兰（*Y. biondii*, KY085894）、武当玉兰（*Y. sprengeri*, JX280401）、玉灯玉兰（*Y. denudata* ‘Lamp’, JN227740）、云南拟单性木兰（*Parakmeria yunnanensis*, KF753638）、宝华玉兰（*Y. zenii*, MH607378）、紫玉兰（*Y. liliiflora*, JX280397）、鹅掌楸（*Liriodendron chinense*, NC_030504）、北美鹅掌楸（*L. tulipifera*, NC_008326）共 11 种木兰科植物的叶绿体基因组数据，其中包括 2 鹅掌楸属植物和其他 9 种木兰科植物。利用 MEGAX 软件，基于邻接法（NJ 法）构建日本厚朴在内的 14 种木兰科植物的系统发育树，观察并分析它们之间的亲缘关系（Yan et al., 2015）。

2 结果与分析

2.1 叶绿体基因组信息

通过 Illumina HiSeq 高通量测序平台测序共获得日本厚朴叶绿体 19 816 708 条原始数据，移除接头和低质量的 Reads，共获得 Clean Reads 19 791 019 条，Q30 为 91.33%。通过在线组装发现日本厚朴全长 160 051 bp，符合目前已知的木兰类植物叶绿体基因组大小范围（159 429~160 183 bp），厚朴叶绿体拥有目前已知木兰类植物的最大基因组（160 183 bp），而凹叶厚朴的叶绿体基因组为 160 099 bp。

三种厚朴叶绿体基因均具有典型的四分区域结构，其中 LSC 分别为 88 210 bp（厚朴）、88 145 bp（凹叶厚朴）和 88 146 bp（日本厚朴），SSC 分别为 18 843、18 832 和 18 771 bp，2 段反向互补重复的 IR 区（IRA 和 IRB）分别为 26 566、26 566 和 26 562 bp。三种厚朴 IR、LSC 和 SSC 区域的 GC 值存在一定的差异，其中 GC 含量最高的区域都为 IR 区，分别为 43.2%、43.1%和 43.2%，LSC 区次之，均为 37.9%，而 SSC 区域的 GC 值最小，仅为 34.2%、34.3%和 34.3%。与其他物种叶绿体基因组相似的是，三种厚朴的 *ycf1* 也跨越了 SSC 和 IRA 区，其中位于 IRA 区域长度均为 1 279 bp，厚朴、凹叶厚朴位于 SSC 区域长度为 4 311 bp，比日本厚朴长 51 bp。除此之外，日本厚朴特有的 *ycf1* 基因还跨越了 SSC 和 IRB 区，其序列长度分别为 29 bp、1 279 bp。

通过基因组结构比较，发现三种厚朴叶绿体的 rRNA 和 tRNA 数量相同（表 1）。在基因数量方面，凹叶厚朴相比厚朴多 2 个单拷贝基因，而两者又较日本厚朴分别少了 4 个和 6 个。在编码区数量方面，凹叶厚朴与厚朴相比多 2 个编码区，而日本厚朴数量最多，比前两者分别多 8 个和 10 个。这揭示了厚朴与凹叶厚朴

的叶绿体基因组可能存在较小的结构和功能差异，而日本厚朴的叶绿体基因组结构和功能可能较其他两者更为丰富。

表 1 三种厚朴叶绿体基因基本信息

Table 1 Basic chloroplast genes information of three *Magnolia* species

名称 Name	全长 Length（bp）	GC(%)	单拷贝基因总数 Total number of single copy genes	编码区 Coding region	rRNA	tRNA
厚朴 *M. officinalis*	160 183	39.2	126	82	8	37
凹叶厚朴 *M. officinalis* subsp. *biloba*	160 099	39.2	128	84	8	37
日本厚朴 *M. hypoleuca*	160 051	39.2	132	92	8	37

2.2 叶绿体内含子、外显子比较

对三种厚朴的叶绿体基因组注释文件进行分析，发现在近百个木兰类叶绿体基因组编码基因中，19 个存在内含子、外显子的差异而 3 个基因的内含子、外显子数量和长度完全一致（表 2）。在这 19 个差异片段中，厚朴与凹叶厚朴的差异主要表现在 11 个，其中 9 个片段的差异区在 I 类内含子（InI，核苷酸数量差异在 20 以内），2 个在 II 类内含子（InII，核苷酸数量差异在 2 以内），还有 1 个差异片段存在 II 类外显子（*rpl2*^{b} 的 EPII，40 bp，此区域也为厚朴、凹叶厚朴叶绿体基因组差异最大的区域）。相较于前两个厚朴物种，日本厚朴存在更多差异，尤其是 *trnI-GAU* 基因的 I 类内含子比前两者多 920 个核苷酸，甚至还多了三种全表达的功能基因（*rps22*^{L}、*ycf1*$^{s\text{-}b}$ 和 *ycf15*^{a,b}）(表 2)，这些结果再次显示厚朴与凹叶厚朴的叶绿体基因组之间的差异较小，而日本厚朴的叶绿体基因组在结构和功能方面较其他两者更为丰富。

在三种厚朴叶绿体的 2 个 *rps12* 基因中，一个正常含有 3 个外显子，而另一个缺失了 I 类外显子，无法正常表达，可能为假基因。植物叶绿体基因普遍具有保守性与突变性（转移或损失等）并存的现象，其转录涉及到复杂的反向剪接等过程，本研究即发现三种厚朴中 *rps12* 的几个外显子为反向排列，且为非顺序式。核糖体小亚基基因的假基因转化在物种进化过程中起到重要作用（Liu et al., 2020），因此 *rps12* 核苷酸序列的差异可为厚朴类亲缘关系研究提供重要参考。三种厚朴叶绿体差异最大的 *ycf3* 基因，其差异区域数达 4 个（I 类、II 类内含子和 I 类、III 类外显子），日本厚朴的 I 类外显子增加了 102 个核苷酸。此外，与木兰类植物的叶绿体基因组条形码 *trnK-UUU*、*trnL* 等相似，本研究针对三种厚朴的比较分析，发现 *trnI-GAU*、*rps16*、*rpoC1*、*clpP*、*ndhA*、*rpl2* 等基因的变异位点率、信息位点率较高，也具备开发为少数木兰植物鉴定条形码的潜力。

表 2 三种厚朴叶绿体基因组内含子、外显子比较

Table 2 Comparison of introns and exons of three *Magnolia* species chloroplast genomes

基因 Gene	厚朴 *M. officinalis*					凹叶厚朴 *M. officinalis* subsp. *biloba*					日本厚朴 *M. hypoleuca*				
	EpI	InI	EpII	InII	EpIII	EPI	InI	EpII	InII	EpIII	EpI	InI	EpII	InII	EpIII
trnK-UUU	37	2498	35			37	2492	35			37	2493	35		
rps16	42	824	246			42	824	246			40	823	221		
atpF	145	709	410			145	707	410			144	709	411		
rpoC1	432	740	1 614			432	734	1 614			432	734	1614		
ycf3	124	733	232	729	151	124	734	232	727	151	226	732	232	727	153
trnL-UAA	35	491	50			35	491	50			35	491	50		
trnV-UAC	39	584	37			39	585	37			39	584	37		
rpl2^{b}	391	661	431			391	661	391			397	661	431		

rps12[b]	114	536	232		26	114	526	232		26	114	536	232		26
clpP	71	786	291	629	246	71	781	291	628	246	71	781	291	628	246
petB	6	784	642			6	784	642			6	792	642		
petD[L]	8	701	475			8	701	475			8	701	475		
rpl16	9	969	399			9	969	399			9	969	399		
ndhB[b]	775	700	758			775	700	758			776	700	755		
trnI-GAU[a]	42	937	35			42	936	35			42	936	35		
trnA-UGC[b]	38	800	35			38	800	35			38	799	35		
ndhA	553	1 082	539			553	1 102	539			552	1 103	540		
trnA-UGC[a]	38	800	35			38	800	35			38	799	35		
trnI-GAU[b]	36	16	36			36	16	36			42	936	35		
ndhB[a]	775	700	758			775	700	758			776	700	755		
rpl2[a]	391	661	431			391	661	431			397	661	431		
trnG-UC/ trnG-GCC	24	770	48			24	767	48			23	768	47		

注：***rps***. 核糖体小亚基；***rpo***. RNA 聚合酶；***ycf***. 开放阅读框；***clp***. 酪蛋白水解蛋白酶；***pet***. 多肽；***ndh***. NADH 脱氢酶；***rpl***. 核糖体大亚基；**Ep**. 外显子；**In**. 内含子；***a***. IRA 区；***b***. IRB 区；***L***. LSC 区。

Note: ***rps***. Ribosomal protein subunit; ***rpo***. RNA polymerase; ***ycf***: Hypothetical chloroplast openreading frame; ***clp***. Caseinolytic protease; ***pet***. polypeptide; ***ndh***. NADH dehydrogenase; ***rpl***. Ribosomal protein large; **Ep**. Expressed region; **In**. Intron; ***a***. IRA region; ***b***. IRB region; ***L***. LSC region.

2.3 叶绿体功能基因比较

在叶绿体四大类基因组成中，厚朴含有光合作用相关基因 46 个、基因表达相关基因 69 个、其他基因 6 个及未知功能基因 5 个。凹叶厚朴的基因表达相关基因少 2 个（*rps3* 和 *rps15*）。而日本厚朴相比于前两者，存在 6 个差异基因，主要集中于基因表达相关基因及未知功能基因，即 *rps3*、*rps15*、*ycf1*、*ycf15*[a,b] 及 *rpl22*（表 3），部分参与多肽形成（Pszczółkowska et al., 2020）功能外，其他基因的作用还存在一定争论，如 *ycf15* 基因在龙葵属（*Amborella*）、萍蓬草属（*Nuphar*）、单子叶和蔷薇类等原始被子植物中是无功能的，甚至在八角属（*Illicium*）、菖蒲属（*Acorus*）、金鱼藻属（*Ceratophyllum*）、毛茛属（*Ranunculus*）等植物的进化过程中已经完全丢失了（Shi et al., 2013），而木兰属（*Magnolia*）和胡椒属（*Piper*）植物基本保留了这类基因。

综合该文内含子与外显子、功能基因的结构分析结果，三种厚朴叶绿体基因组的差异主要分布于 LSC 区和 IR 区，涉及核糖体大亚基、核糖体小亚基和未知功能基因类群，尤其是诸多未知功能的基因，而木兰属中这类疑似非功能基因的结构特征、变异原因，以及与适生环境差异的关联性，有待进一步深入研究。

表 3 三种厚朴叶绿体的功能基因比较

Table 3 Comparison of chloroplast functional genes of three *Magnolia* species

基因功能 Gene function	基因种类 Gene type	日本厚朴 *M. hypoleuca*	厚朴 *M. officinalis*	凹叶厚朴 *M. officinalis* subsp. *biloba*
光合作用 Photosynthesis	ATP 合酶亚基 ATP synthase subunit	*atpA, atpB, atpE, atpF, atpH, atpI*	*atpA, atpB, atpE, atpF, atpH, atpI*	*atpA, atpB, atpE, atpF, atpH, atpI*
	光合系统I亚基 Photosynthetic system I subunit	*psaA, psaB, psaC, psaI, psaJ*	*psaA, psaB, psaC, psaI, psaJ*	*psaA, psaB, psaC, psaI, psaJ*
	光合系统II亚基 Photosynthetic system II subunit	*psbA, psbB, psbC*[L], *psbC*[L], *psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbT, psbZ, psbH*	*psbA, psbB, psbC*[L], *psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, psbH*	*psbA, psbB, psbC*[L], *psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, psbH*

	NADH 脱氢酶亚基 NADH dehydrogenase subunit	*ndhA*, *ndhB*[a], *ndhB*[b], *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*[L], *ndhK*[L]	*ndhA*, *ndhB*[a], *ndhB*[b], *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*	*ndhA*, *ndhB*[a], *ndhB*[b], *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*
	细胞色素 b/f 复合体 Cytochrome b/f complex	*petA*, *petB*, *petD*[L], *petD*[L], *petG*, *petL*, *petN*	*petA*, *petB*, *petD*[L], *petG*, *petL*, *petN*	*petA*, *petB*, *petD*[L], *petG*, *petL*, *petN*
基因表达 Gene expression	二磷酸核酮糖羧化酶亚基 Ribulose diphosphate carboxylase subunit	*rbcL*	*rbcL*	*rbcL*
	核糖体大亚基 Large ribosomal subunit	*rpl14*, *rpl16*, *rpl2*[a], *rpl2*[b], *rpl20*, *rpl22*, *rpl23*[a], *rpl23*[b], *rpl32*, *rpl33*, *rpl36*	*rpl14*, *rpl16*, *rpl2*[a], *rpl2*[b], *rpl20*, *rpl22*, *rpl23*[a], *rpl23*[b], *rpl32*, *rpl33*, *rpl36*	*rpl14*, *rpl16*, *rpl2*[a], *rpl2*[b], *rpl20*, *rpl22*, *rpl23*[a], *rpl23*[b], *rpl32*, *rpl33*, *rpl36*
	核糖体小亚基 Small ribosomal subunit	*rps11*, *rps12*[a], *rps12*[b], *rps12*[L], *rps14*, *rps15*, *rps16*, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*[a], *rps7*[b], *rps8*	*rps11*, *rps12*[a], *rps12*[b], *rps14*, *rps16*, *rps18*, *rps19*, *rps2*, *rps4*, *rps7*[a], *rps7*[b], *rps8*	*rps11*, *rps12*[a], *rps12*[b], *rps14*, *rps15*, *rps16*, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*[a], *rps7*[b], *rps8*
	依赖于 DNA 的 RNA 聚合酶亚基 DNA-dependent RNA polymerase subunit	*rpoA*, *rpoB*, *rpoC1*, *rpoC2*[L], *rpoC2*[L]	*rpoA*, *rpoB*, *rpoC1*, *rpoC2*	*rpoA*, *rpoB*, *rpoC1*, *rpoC2*
	核糖体 RNA Ribosomal RNAs	*rrn16*[a], *rrn23*[a], *rrn4.5*[a], *rrn5*[a], *rrn5*[b], *rrn23*[b], *rrn16*[b], *rrn4.5*[b]	*rrn16*[a], *rrn23*[a], *rrn23*[a], *rrn4.5*[a], *rrn5*[a], *rrn5*[b], *rrn23*[b], *rrn23*[b], *rrn16*[b], *rrn4.5*[b]	*rrn5*[b], *rrn23*[b], *rrn23*[b], *rrn16*[b], *rrn4.5*[b], *rrn16*[a], *rrn23*[a], *rrn23*[a], *rrn4.5*[a], *rrn5*[a],
	乙酰辅酶 A 羧化酶亚基 Acetyl-CoA carboxylase subunit	*accD*	*accD*	*accD*
	C 型细胞色素合成酶 C-type cytochrome synthase	*ccsA*	*ccsA*	*ccsA*
其他基因 Other genes	膜蛋白 Membrane protein	*cemA*	*cemA*	*cemA*
	蛋白酶 Protease	*clpP*	*clpP*	*clpP*
	翻译起始因子 Translation initiation factor	*infA*	*infA*	*infA*
	成熟酶	*matK*	*matK*	*matK*

	Mature enzyme			
未知功能基因 Unknown functional gene	保守的开放阅读框 Conservative open reading frame	*ycf1*$^{s\text{-}a}$, *ycf1* $^{s\text{-}b}$, *ycf2*^{a}, *ycf2*^{b}, *ycf3*, *ycf4*, *ycf15*^{a}, *ycf15*^{b}	*ycf1* $^{s\text{-}a}$, *ycf2*^{a}, *ycf2*^{b}, *ycf3*, *ycf4*	*ycf1* $^{s\text{-}a}$, *ycf2*^{a}, *ycf2*^{b}, *ycf3*, *ycf4*

注：***psa***. 光合器；***acc***. 乙酰辅酶 A 羧化酶；***rbc***. 二磷酸核酮糖羧化酶；***ccs***. C 型细胞色素合成基因；***cem***. 叶绿体包膜蛋白；***mat***. 成熟酶基因；***a***. IRA 区；***b***. IRB 区；***s-a***. 横跨 SSR 区和 IRA 区；***s-b***. 横跨 SSR 区和 IRB 区。

Note: ***psa***. Photosynthetic apparatus; ***acc***. Acetyl-coa carboxylase; ***rbc***. Ribulose bisphosphate carboxylase; ***ccs***. TypeC cytochrome synthesis gene; ***cem***. Chloroplast envelope membrane protein; ***mat***. Mature enzyme gene; ***a***. l IRA region; ***b***. IRB region; ***s-a***. Across the SSR and IRA regions; ***s-b***. Across the SSR and IRB regions.

2. 4 部分木兰科植物的亲缘关系

由图 1 可知，在 14 种近缘物种叶绿体基因组构建的 NJ 发育树（节点支持率均在 75%以上）中，鹅掌楸属和其他属是木兰科中的两个独立的谱系，这与之前报道的结果一致（陈凯，2019）。在木兰属的 2 个次级类群中，三种厚朴、天女木兰、荷花玉兰聚类在一起，其中日本厚朴与凹叶厚朴植物关系最为密切，其次是厚朴。木兰属中另外一个次级类群包括云南拟单性木兰、玉灯玉兰、紫玉兰、武当玉兰、望春玉兰、星华玉兰、宝华玉兰，它们与三种厚朴亲缘关系由近及远。

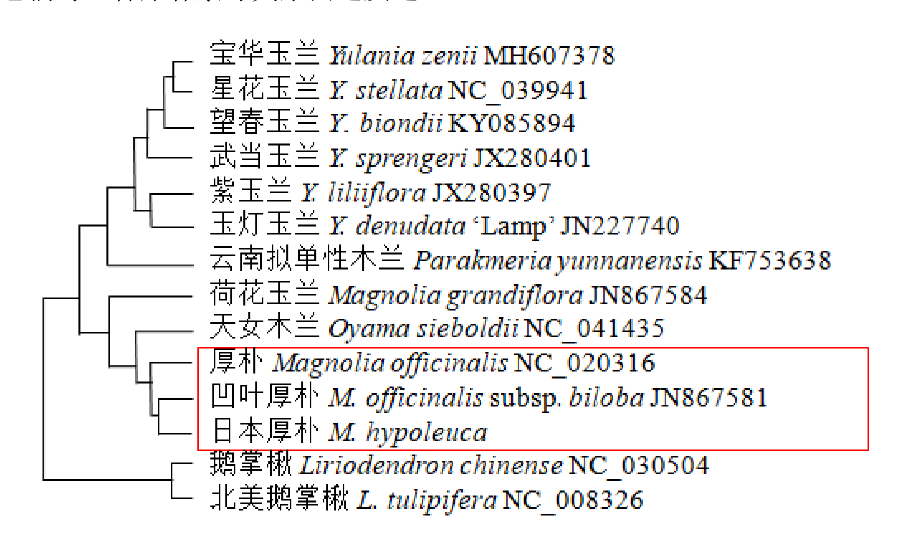


图 1 部分木兰科物种系统进化树

Fig.1 Phylogenetic tree of partial Magnoliaceae plants

3 讨论与结论

本研究使用全基因组测序来组装和比较三种厚朴的叶绿体基因组，以此来推测三种厚朴之间多种生物学性状差异的形成原因。低温环境下，植物最先受到抑制的生理代谢过程是光合作用，低温能显著降低光合效率、CO_2 同化作用和光系统II活性，从而影响植物的正常生长发育（王璐等，2020）。通过三种厚朴叶绿体的功能基因比较，发现日本厚朴 *psbC* 基因在数量上较另两者增加了一个，这可能缘于 *psbC* 基因的自我复制。*psbC* 和 *psbD* 是光合系统II的关键基因，例如小麦成熟叶绿体 psbC 蛋白的合成发生于含 *psbD* 序列的转录本上，形成 *psbD-psbC* 复合物，它们的转录水平受到光诱导，通过增强光合系统II亚基的合成和维持能力，提高叶绿体的合成速度（Gamble et al., 1988; Gamble & Mullet, 1989），从而提高叶绿体的光合作用强度，进而加快植物的生长发育。此外，*psbD-psbC* 基因协同转录的 mRNAs 可以翻译产生 D2 和 CP43 蛋白，与反应中心 D1、CP47、放氧复合体蛋白及捕光复合体 II 等蛋白，共同参与叶绿体光合系统 II 的光合电子传递，在强光条件下维持光合系统 II 的功能（庄焜扬，2020）。*psbD-psbC* 复合体直接参与光合系统II亚基的形成，在生长旺盛和存在颜色差异的叶片中发挥着保护光合系统、减少强光伤害的作用（Adachi et al., 2012），这在花叶矢竹叶等光合系统II中 *psbD* 基因的研究中也得到证实（许冰清等，2015）。因此，日本厚朴叶绿体增加的 *psbC* 基因表明其具备合成更多 *psbD-psbC* 复合体与 D2 蛋白的潜力，这可能是日本厚朴生长速度更快、适应较北纬度地区低温和强光胁迫环境的原因之一。

在高变异位点率的功能基因中，*rpl22* 为日本厚朴中特有，其位于 LSC 区，属于基因表达相关功能基因，

可用于物种鉴定（Feng et al., 2019）。除此外，本研究发现三种厚朴的差异基因主要分布于 *rpl22* 所属的基因簇中，它们组成 *rpl23* 操纵子的大转录单元，翻译为核糖体大蛋白亚基。该研究发现，此操纵子在三种厚朴中存在显著的结构差异：*rp123-rp12-rps19-rp122*（日本厚朴）*-rps3*（凹叶厚朴、日本厚朴）*-rp116-rp114-rps8-infA-rpl36-rps11-rpoA*。通过三个物种的亲缘关系可知，凹叶厚朴、日本厚朴与厚朴分化的过程中，该操纵子 *rps19-rp116* 之间原有的核苷酸空隙逐次被 *rp13*、*rp122* 基因填充，同时造成操纵子中多个功能基因的内含子、外显子核苷酸发生了增减（减少为主），功能蛋白的数量和活性也发生了改变。植物叶绿体除了光合作用外，还部分参与氨基酸、核苷酸、脂类和淀粉等各类成分的生物合成，支撑起该植物的生物学性状（Namgung et al., 2021），日本厚朴具备更丰富的核糖体基因簇，这无疑形成了其更丰富的生物学特性。

tRNA 作为核酸信息水平和蛋白质功能水平的适配器，在蛋白质翻译中起着核心作用，其结构修饰影响植物体的温度适应能力（Lorenz et al., 2017）。本研究发现日本厚朴 tRNA-ALA 比其他两种厚朴多 3 个，具有快速转运和积累丙氨酸的潜力，而游离态的丙氨酸能抵抗多种外界环境刺激（Mustroph et al., 2014），如寒冷因子。除了数量因素外，三种厚朴叶绿体中 tRNA 对应的氨基酸种类也不同。日本厚朴为本属植物中分布较为北端（如千岛群岛）的物种，其生境中弱光和低温信号交叉影响着植物的生长发育，在这种光温胁迫条件下，tRNA 可能一方面积累更多的游离丙氨酸，另一方面降低正常蛋白质翻译速率，将正常蛋白质合成转向应激反应蛋白质的合成（冯德江等，2002），最终使得植株具备更强的耐寒性，这又为厚朴优良品种的分子筛选和培育提出了参考。

参考文献：

ADACHI Y, KURODA H, YUKAWA Y, et al., 2011. Translation of partially overlapping *psbD-psbC* mRNAs in chloroplasts: the role of 5'-processing and translational coupling [J]. Nucl Acid Res, 40(7): 3152-8.

CHEN K, 2019. Study on the structural variation of chloroplast genome in Magnoliaceae and screening of hypervariable genes[D]. Hangzhou: Zhejiang A & F University: 7-15. [陈凯，2019. 木兰科植物叶绿体基因组结构变异研究及高变基因的筛选[D]. 杭州：浙江农林大学：7-15.]

FENG DJ, LIU X, LI XG, et al., 2002. Relationship between tRNA abundance and gene expression[J]. Chin J Biol Eng, 6: 4-8. [冯德江，刘翔，李旭刚，等，2002. tRNA 丰度与基因表达的关系[J]. 中国生物工程杂志，6:4-8.]

FENG Z, ZHANG L, WU Y, et al., 2019. The *Rpf84* gene, encoding a ribosomal large subunit protein, RPL22, regulates symbiotic nodulation in *Robinia pseudoacacia* [J]. Planta, 250(6): 1897-1910.

GAMBLE PE, SEXTON TB, MULLET JE, 1988. Light-dependent changes in *psbD* and *psbC* transcripts of barley chloroplasts: accumulation of two transcripts maintains *psbD* and *psbC* translation capability in mature chloroplasts [J]. EMBO J, 7(5): 1289-1297.

GAMBLE PE, MULLET JE, 1989. Blue light regulates the accumulation of two *psbD-psbC* transcripts in barley chloroplasts [J]. EMBO J, 8(10): 2785-94.

KWON OJ, OH CH, 2015. Naturalization of landscaping woody plant, *Magnolia obovata* potentially invasive species [J]. J M Sci, 12(1): 30-38.

LIU SS, WANG Z, WANG H, et al., 2020. Patterns and rates of plastid *rps12* gene evolution inferred in a phylogenetic context using plastomic data of ferns [J]. Sci Rep, 10(1): 9394.

LORENZ C, LUNSE CE, MORL M, 2017. tRNA modifications: Impact on structure and thermal adaptation [J]. Biomolecules, 7(2): 35.

MATTHES N, WESTPHAL K, HALDEMANN C, et al., 2020. Validation of a modified CTAB method for DNA extraction from protein-rich maize feedstuffs [J]. J Consum Prot Food Saf, 15(4): 331-340.

MUSTROPH A, BARDING GA, KAISER KA, et al., 2014. Characterization of distinct root and shoot responses to low-oxygen stress in *Arabidopsis* with a focus on primary C and N-metabolism [J]. Plant cell Environ, 37(10): 2366-2380.

NAMGUNG J, DO HDK, KIM C, et al., 2021. Complete chloroplast genomes shed light on phylogenetic relationships, divergence time, and biogeography of *Allioideae* (Amaryllidaceae) [J]. Sci Rep, 11(1): 3262.

OGUCHI R, HIURA T, HIKOSAKA K, 2017. The effect of interspecific variation in photosynthetic plasticity on

4-year growth rate and 8-year survival of understorey tree seedlings in response to gap formations in a cool-temperate deciduous forest [J]. Tree Physiol, 37(8): 1113-1127.

PENG MJ, 2020. Resources and genetic analysis of *Magnolia* in Baotianmanshan district of Nanzhao[D]. Zhengzhou: Henan Agricultural University: 2-40. [彭梦婕，2020. 南召境内宝天曼山区木兰属植物资源概况及遗传分析[D]. 郑州：河南农业大学：2-40.]

PSZCZOLKOWSKA A, ANDROSIUK P, JASTRZEBSKI JP, et al., 2020. *rps3* as a candidate mitochondrial gene for the molecular identification of species from the *Colletotrichum acutatum* species complex [J]. Genes (Basel), 11(5): 552.

SHI C, LIU Y, HUANG H, et al., 2013. Contradiction between plastid gene transcription and function due to complex posttranscriptional splicing: an exemplary study of *ycf15* function and evolution in angiosperms [J]. PLoS ONE, 8(3): e59620.

WANG L, LI YL, XIONG H, et al., 2020. Effects of temperature stress on leaf structure and photosynthetic characteristics of *Castanea henryi* seedlings[J]. J Jiangxi Agric, 42(4): 692-699. [王璐, 李艳丽, 熊欢, 等，2020. 温度胁迫对锥栗幼苗叶片结构及光合特性的影响[J]. 江西农业大学学报, 42(4)：692-699.]

XU AL, XIAO GL, BI XL, et al., 2021. Rapid analysis of chemical constituents of *Magnolia officinalis* Wenzhong decoction[J]. N Chin Mat Med Clin Pharm, 32(2): 252-258. [胥爱丽, 肖观林, 毕晓黎, 等，2021. 厚朴温中汤化学成分快速分析[J]. 中药新药与临床药理，32(2)：252-258.]

XU BQ, AN MM, JIANG K, et al., 2015. Cloning and functional analysis of the chloroplast *psbD* gene of *Phyllostachys japonicus*[J]. J Zhejiang Agric For, 32(4): 557-565. [许冰清, 安苗苗, 姜可以, 等，2015. 花叶矢竹叶绿体 *psbD* 基因的克隆与功能分析[J]. 浙江农林大学学报，32(4)：557-565.]

YAN A, LAI XJ, LI XD, et al., 2015. Analyses of the complete genome and gene expression of chloroplast of sweet potato [J]. PLoS ONE, 10(4): 1-25.

ZHAO Q, YU JX, QIN YW, et al., 2021. Assembly and sequence analysis of chloroplast genomes in rapeseed based on high-throughput sequencing[J]. Chin Tradit Herbal Med, 52(6): 1744-1750. [赵祺, 余佳兴, 秦宇雯， 等，2021. 基于高通量测序的菜头肾叶绿体基因组的组装及序列分析[J]. 中草药，52(6)：1744-1750.]

ZHUANG KY, 2020. The WHIRLY1 protein in tomato *WHIRLY1* is a bilocentric protein and a bilocentric protein under temperature stress[D]. Tai'an: Shandong Agricultural University: 6-20. [庄焜扬，2020. 温度胁迫下番茄叶绿体与细胞核双定位 WHIRLY1 蛋白的功能分析[D]. 泰安：山东农业大学：6-20.]